

## SEQUENCE LISTING

<110> INVITROGEN CORPORATION  
DALBY, Brian  
BENNETT, Robert

<120> DELIVERY OF FUNCTIONAL PROTEIN SEQUENCES  
BY TRANSLOCATING POLYPEPTIDES

<130> INVIT1280-1

<150> PCT/US00/08571  
<151> 2000-03-31

<150> 60/127,467  
<151> 1999-03-31

<160> 21

<170> FastSEQ for Windows Version 4.0

<210> 1  
<211> 6404  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> vector pVP22/Myc-His

<400> 1  
gacggatcgg gagatctccc gatcccstat ggtcactct cagtaatac tgctctgatg 60  
ccgcatagtt aagccagtat ctgctccctg cttgtgttt ggaggtcgct gagtagtgcg 120  
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180  
ttagggtagt gcgccccggc ctgcttcggc atgtacgggc cagatatacg cgttgcattt 240  
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300  
tggagttccg cggtacataa cttacggtaa atggccgcgc tggctgaccg cccaacgacc 360  
cccgccccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420  
attgacgtca atgggtggac tatttacggt aaactgccc cttggcagta catcaagtgt 480  
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540  
atgcccagta catgaccta tgggacttgc ctacttggca gtacatctac gtattagtca 600  
tcgcttattac catggtgatg cgggtttggc agtacatcaa tgggctgttga tagcggtttg 660  
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttgc tttggcacc 720  
aaaatcaacg ggacttcca aatgtcgtt acaactccgc cccattgacg caaatggcg 780  
gtaggcgtgt acgggtggag gtctatataa gcagagctct ctggcttaact agagaaccca 840  
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggcttagt 900  
taagcttatt atgacctctc gccgctccgt gaagtcgggt ccgcggggagg ttccgcgcga 960  
tgagtacgag gatctgtact acaccccttc tttaggtatg gcgagtcccg atagtccgc 1020  
tgacacctcc cggccgtggcg ccctacagac acgttcgcgc cagagggcg agtccgttt 1080  
cggtccgtac gacgagtcgg attatgcctt ctacggggc tcgtcttccg aagacgacga 1140  
acacccggag gtccccggc cgcggcggtcc cgtttccggg gcggtttgt ccggccgggg 1200  
gcctgcgcgg ggcgcctccgc caccggctgg gtccggaggg gccggacgc cacccaccac 1260  
cgccccccgg gccccccggaa cccagcggtt ggcgtctaag gccccccggg ccccgccggc 1320  
ggagaccacc cgcggcagga aatcgccca gccagaatcc gccgcactcc cagacgcccc 1380  
cgcgctcgacg ggcggcaaccc gatccaagac accccgcgcag gggctggcca gaaagctgca 1440  
ctttagcacc gcccccccaa accccgacgc gccatggacc ccccggtgg ccggctttaa 1500  
caagcgctc ttctgcgcgg cggtcggggcg cctggcggcc atgcatgccc ggatggcg 1560  
tgtccagctc tgggacatgt cgcgtccggc cacagacgaa gacctcaacg aactccttgg 1620  
catcaccacc atcccggtga cggtctgcga gggcaaaaac ctgcttcagc ggcggcaacga 1680  
gttggtgaat ccagacgtgg tgcaggacgt cgacgcggcc acggcgactc gagggcggttc 1740

tgcggcgtcg	cgc	ccccaccg	agcgacctcg	agccccagcc	cgctccgctt	ctcgccccag	1800	
acggcccgtc	gagg	gtaccg	agctcgatc	cactagtcca	gtgtggtgg	attctgcaga	1860	
tatccagcac	agtgg	ccggcc	gctcgagtct	agagggccc	cggttcaac	aaaaactcat	1920	
ctcagaagag	gatct	gaata	tgcataccgg	tcatcatcac	catcaccatt	gagttaaac	1980	
ccgctgatca	gcctcgact	g	tcgccttctag	ttgccagcca	tctgtgttt	gccctcccc	2040	
cgtgccttcc	ttgac	ccttgg	aagg	tgccac	tccactgtc	cttcctaataa	2100	
aattgcac	cattgt	cgtc	gttaggtgtca	ttctattctg	gggggtgggg	tggggcagga	2160	
cagcaagggg	gagg	atttggg	aagacaatag	caggcatgt	ggggatgcgg	tgggctctat	2220	
ggcttcttag	gcgg	aaaagaa	ccagctgggg	ctctaggggg	tatccccac	cgcctgttag	2280	
cggcgcatta	agcgc	gggg	gtgtgggtt	tacgcgcage	gtgaccgcta	cacttgccag	2340	
cgcgcctagcg	cccgc	tcctt	tcgccttctt	cccttc	ctgcacgt	tcgcggctt	2400	
tccccgtca	gctctaaatc	gggg	catccc	tttaggttc	cgat	ttagtgcata	2460	
cctcgacccc	aaaaaact	atagg	gtgtgttgc	tgggtcacgt	agtggccat	cgcctgtata	2520	
gacggtttt	cgcc	ccccc	tttgc	cggtggatc	cacgttctt	aatagtggac	2580	
aactggaaca	acact	caacc	ctatctcggt	ctattcttt	gatttataa	aaatthaac	2640	
gatttccggc	tatt	gtttaa	aaaatgagct	gatttacaa	aaatthaac	cgaattaatt	2700	
ctgtggaat	tgtgt	cagtt	agggtgtgg	aagtcccc	gctcccccagg	caggcagaag	2760	
tatgcaaagc	atgc	atctca	attagtca	aaccagggt	ggaaagtccc	caggctcccc	2820	
agcaggcaga	agtat	gcaaa	gcatgcatt	caattagtca	gcaaccatag	tccgc	2880	
aactccgccc	atccc	cccc	taactccg	cagtccg	cattctccgc	ccatggctg	2940	
actaattttt	tttattt	at	tgatc	cagaggccg	gcctctgagc	tattccagaa	3000	
gtagtgagga	ggc	ttttt	tttgc	ggggcttgg	aagctccc	gagcttgat	3060	
atccat	tttc	ggatctgatc	aagagacagg	atgaggatcg	tttgc	ttgaacaaga	3120	
tggattgcac	gcag	gttctc	cgcccg	gggtggaggg	ctattcg	atgactggc	3180	
acaacagaca	atcg	gctgt	ctgat	ccgc	ctgtcagc	agggcgccc	3240	
ggttctttt	gtca	agaccg	acctgt	cccg	gactgc	acgaggcagc	3300	
gcggctatcg	tgg	ctggcc	cgac	gggg	gtgtgc	acgttgtcac	3360	
tgaagcggga	agg	gactggc	tgct	atttgg	ggcaggatc	tcctgtcatc	3420	
tcaccttgct	cct	gcgaga	aagt	atccat	catggct	gcaatgc	3480	
gctt	gatcc	gctact	gc	cattcgac	ccaa	catcgatcg	3540	
tactcgat	gaag	ccgg	gt	cgatct	gacg	aaag	atcaggggct	3600
cgcgc	ccat	cgat	cc	aggctca	ggcgc	ccgcacgg	3660	
cgtgacccat	ggc	gatgc	ct	gttgc	ccat	gttgcata	3720	
attcatcgac	tgt	ggcc	gg	ccgt	gaa	catcgatcg	3780	
ccgtgatatt	gct	gaag	ag	atgg	act	cgatcgatcg	3840	
tatcgccgct	ccc	gatt	tcg	atcg	tttgc	ttgttacgg	3900	
agcgggactc	tgg	gtt	ccg	aaat	atcg	atcgatcg	3960	
atttcgattc	cac	ccg	cc	ttctat	gat	ccatcacgag	4020	
ccggcttggat	gat	cct	cc	gggtgg	atc	atcgatcg	4080	
tgtttattgc	ag	ctt	at	tttgc	atc	atcgatcg	4140	
aagcattttt	ttc	act	tg	tttgc	act	atcgatcg	4200	
atgtctgtat	acc	gtc	gac	tct	atc	atcgatcg	4260	
ctgtgtaaa	ttg	ttat	cc	gttgc	atc	atcgatcg	4320	
gtaaagcctg	ggg	gtt	cc	tttgc	atc	atcgatcg	4380	
ccgc	ttt	c	tc	tttgc	atc	atcgatcg	4440	
ggagaggcgg	tttgc	tatt	ggg	ccgtt	atc	atcgatcg	4500	
cgtcg	ttcg	gtc	ggc	gttgc	atc	atcgatcg	4560	
cagaatcagg	ggat	aaac	gc	tttgc	atc	atcgatcg	4620	
accgtaaaaa	ggcc	cg	tttgc	tttgc	atc	atcgatcg	4680	
acaaaaatcg	acg	ctt	cc	tttgc	atc	atcgatcg	4740	
cgttcccc	tgg	aaat	cc	tttgc	atc	atcgatcg	4800	
acctgtccgc	ttt	cc	tc	tttgc	atc	atcgatcg	4860	
atctcagttc	gg	tttgc	tttgc	tttgc	atc	atcgatcg	4920	
agcccgacc	ctg	cc	tttgc	tttgc	atc	atcgatcg	4980	
acttatcgcc	act	gg	cc	tttgc	atc	atcgatcg	5040	
gtgctacaga	gtt	tttgc	tttgc	tttgc	atc	atcgatcg	5100	
gtatctgcgc	tct	gtt	cc	tttgc	atc	atcgatcg	5160	
gcaaacaac	cac	cc	tttgc	tttgc	atc	atcgatcg	5220	
aaaaaaaagg	atc	taa	gg	tttgc	atc	atcgatcg	5280	
acaaaaactc	acg	tttgc	tttgc	tttgc	atc	atcgatcg	5340	
tccttttaaa	ttaa	aaat	tttgc	tttgc	atc	atcgatcg	5400	
			aaatctaa	tttgc	atc	atcgatcg		

2020-04-26 09:26:56

ctgacagtta	ccaatgctta	atcagtggagg	caccttatctc	agcgatctgt	ctatttcgtt	5460
catccatagt	tgcctgactc	cccgtcggt	agataactac	gatacgggag	ggcttaccat	5520
ctggccccag	tgctgaatg	ataccgcgag	acccacgctc	accggctcca	gatttatcag	5580
caataaacca	gccagccgga	agggccgagc	gcagaagtgg	tcctgcaact	ttatccgcct	5640
ccatccagtc	tattaattgt	tgccgggaag	ctagagtaag	tagttcgcca	gttaatagtt	5700
tgcgcaacgt	tgttgcatt	gtcacagggc	tcgtgggtgc	acgctcgctc	tttggatgg	5760
cttcattcag	ctcccggtcc	caacgatcaa	ggcgagttac	atgatcccc	atgttgtca	5820
aaaaagcggt	tagctccccc	ggtcctccga	tcgtgtcag	aagtaagttg	gccgcagtgt	5880
tatcaactcat	ggttatggca	gcaactgcata	atttcattac	tgtcatgcca	tccgtaagat	5940
gctttctgt	gactggtgag	tactcaacca	agtcaattctg	agaatagtgt	atgcggcgac	6000
cgagttgctc	ttgcccggcg	tcaatacggg	ataataccgc	gccacatagc	agaactttaa	6060
aagtgcgtat	cattggaaaa	cgttcttcgg	ggcgaaaact	ctcaaggatc	ttaccgctgt	6120
ttagatccag	ttcgatgtaa	cccactcggt	cacccaactg	atcttcagca	tctttactt	6180
tcaccagcgt	ttctgggtga	gcaaaaaacag	gaaggcaaaa	tgccgcaaaa	aagggaataa	6240
ggcgacacg	gaaatgttga	atactcatac	tcttccttt	tcaatattat	tgaagcattt	6300
atcagggtta	ttgtctcatg	agcggataca	tatttgaatg	tatttagaaa	aataaacaaa	6360
taggggttcc	gcbcacattt	ccccgaaaag	tgccacactga	cgtc		6404

<210> 2  
<211> 6420

<212> DNA

<213> Artificial Sequence

<220>  
<223> vector pVP22/Myc-His-TOPO

<400> 2						
gacggatcgg	gagatctccc	gatcccstat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatatgtt	aagccagtat	ctgctccctg	cttgtgttt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatt	aagaatctgc	180
ttagggtag	gcgtttcg	ctgcttcg	atgtacgggc	cagatatacg	cggtgacatt	240
gattattgac	tagtattaa	tagaatcaa	ttacggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacgtaa	atggccgc	tggctgaccg	cccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacgt	aaactgccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgcccagta	catgaccta	tgggactt	ctacttggca	gtacatctac	gtattagtca	600
tcgcttattac	catggatgt	cgggtttggc	agtacatcaa	tgggcgtgga	tagcggttt	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtca	tgggagttt	tttggcacc	720
aaaatcaacg	ggacttcca	aatgtcgt	acaactccgc	cccatggacg	caaatggcgc	780
gtaggcgtgt	acggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcaactatag	ggagacccaa	gctggctagt	900
taagcttatt	atgacctctc	gccgctccgt	gaagtcgggt	ccgcgggagg	ttccgcgcga	960
tgagtacgag	gatctgtact	acacccctc	ttcaggatg	gcgagttcc	atagtccgc	1020
tgacacctcc	cggcgtggc	ccctacagac	acgctcgcc	cagagggc	aggccgtt	1080
cgtccagttac	gacgagtcgg	attatgcct	ctacggggc	tcgtcttcc	aagacgacga	1140
acacccggag	gtccccccg	cgcggcg	cgttccgg	gcgggttt	ccggccgg	1200
gcctgcgcgg	gcgcctccgc	cacccgctgg	gtccggaggg	gcggacgca	cacccaccac	1260
cggcccccgg	gccccccgaa	cccagcggt	ggcactaag	gccccccg	ccccggcgg	1320
ggagaccacc	cgcggcagga	aatcgccca	gccagaatcc	gcgcactcc	cagacgccc	1380
cgcgtcgacg	gcgcacaccc	gatccaagac	acccgcgc	ggcgtggcca	gaaagctgca	1440
ctttagcacc	gcccccccaa	accccgacgc	gccatggacc	ccccgggtgg	ccggctttaa	1500
caagcgcgtc	ttctgcgccc	cggtcggggc	cctggcggcc	atgcatgccc	ggatggcggc	1560
ggtccagctc	tgggacatgt	cgcgtcccg	cacagacgaa	gacctaact	aactccttgg	1620
catcaccacc	atccgcgt	cggctcg	ggccaaaaac	ctgcttc	cgcccaacga	1680
gttggtgaat	ccagacgtgg	tgcaggacgt	cgacgcggcc	acggcactc	gagggcgttc	1740
tgccggcgtcg	cggcccccac	agcgcac	agccccagcc	cgctccgctt	ctcgccccag	1800
acggcccg	gagggttacc	agctcgatc	cactagtca	gtgtggtgg	attgcctt	1860
aggcattc	tgcagatatc	cagcacagtg	gcggccgctc	gagtctagag	ggcccgcggt	1920
tcaaacaaa	actcatctca	gaagaggatc	tgaatatgca	taccggat	catcaccatc	1980
accattgagt	ttaaacccgc	tgatcagc	cgactgtgc	cgccatctg		2040

ttgtttgcc	ctccccgtg	cttccttga	cccttggaaagg	tgccactccc	actgtcctt	2100
cctaataaaa	tgagggaaatt	gcategcatt	gtctgagtag	gtgtcattct	attctggggg	2160
gtgggggtgg	gcaggacagc	aaggggggagg	attgggaaga	caatagcagg	catgctgggg	2220
atgcgggtgg	ctctatggct	tctgaggcgg	aaagaaccag	ctggggctct	agggggatc	2280
cccacgcgc	ctgttagcggc	gcattaagcg	cggcgggtgt	ggtggttacg	cgagcgtga	2340
ccgctacact	tgccagcggc	ctagcggccg	ctcccttcge	tttcttccct	tcctttctcg	2400
ccacgttcgc	cggcttcccc	cgtaagcgc	taaatcgggg	catccctta	gggttccgat	2460
ttagtgtctt	acggcacctc	gaccccaaaa	aacttgatta	gggtgatggt	tcacgtatg	2520
ggccatcgcc	ctgatagacg	gttttgcggc	ctttgacgtt	ggagtccacg	ttctttaata	2580
gtggactctt	gttccaaact	ggaacaacac	tcaaccctat	ctcggtctat	tcttttatt	2640
tataagggt	tttggggatt	tcggcctatt	ggttaaaaaaa	tgagctgatt	taacaaaat	2700
ttaacgcgaa	ttaattctgt	ggaatgtgtg	tcagtttaggg	tgtggaaagt	ccccaggctc	2760
cccaggcagg	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	aggtgtggaa	2820
agtccccagg	ctccccagca	gdcagaagta	tgcaaaagcat	gcatctcaat	tagtcagcaa	2880
ccatagtccc	gccccctaact	ccgccccatcc	cgccccctaac	tccgcccagt	tccgcccatt	2940
ctccgcccc	tggctgacta	attttttta	tttatgcaga	ggccgaggcc	gcctctgcct	3000
ctgagctatt	ccagaagtag	tgaggaggt	tttttggagg	cctaggctt	tgaaaaaagc	3060
tcccgggagc	ttgtatatcc	attttcggat	ctgatcaaga	gacaggatga	ggatcgttc	3120
gcatgattga	acaagatgga	ttgcacgcag	gttctccgc	cgcttgggt	gagaggctat	3180
tcggctatga	ctgggcacaa	cagacaatcg	gctgctctga	tgccgcccgt	ttccggctgt	3240
cagcgcaggg	gcccgggtt	ctttttgtca	agaccgacct	gtccggtgcc	ctgaatgaac	3300
tgcaggacga	ggcagcgcgg	ctatcgtggc	tggccacgac	gggcgttcc	tgcgcagctg	3360
tgctcgacgt	tgtcaactgaa	gccccggagg	actggctgt	attggcgaa	gtgccggggc	3420
aggatctcct	gtcatctcac	cttgcctcctg	ccgagaaaagt	atccatcatg	gctgatgcaa	3480
tgcggcggct	gcatacgctt	gatccggcta	cctgcccatt	cgaccaccaa	gcaaaacatc	3540
gcatcgagcg	agcacgtact	cgatggaaag	ccggcttctgt	cgatcaggat	gatctggacg	3600
aagagcatca	ggggctcgcg	ccagccgaac	tgttcgccag	gctcaaggcg	cgcatgccc	3660
acggcgagga	tctcgtcgt	acccatggcg	atgcctgctt	gccgaatatc	atggtgaaa	3720
atggccgctt	ttctgatttcc	atcgactgt	gccggctggg	tgtggcggac	cgctatcagg	3780
acatagcggt	ggctaccctgt	gatattgctg	aagagcttgg	cgcgaaatgg	gctgaccgct	3840
tcctcgtgt	ttacgtatc	gcccgtcccg	attcgcagcg	catcgcccttc	tatcgcccttc	3900
ttgacgagtt	cttctgagcg	gactctgtgg	gttcgcgaaa	tgaccgacca	agcgacgccc	3960
aacctgccc	cacgagattt	cgattccacc	gccgccttct	atgaaagggtt	gggcttcgg	4020
atcgtttcc	gggacccgg	ctggatgatc	ctccagcgcg	gggatctcat	gctggagttc	4080
ttcgcacc	ccaacttgtt	tattgcagct	tataatggtt	acaaataaaag	caatagcatc	4140
acaaatttca	caaataaaagc	attttttca	ctgcattcta	gttgtggttt	gtccaaactc	4200
atcaatgtat	cttacatgt	ctgtataccg	tcgacctcta	gctagagctt	ggcgtaatca	4260
tggcatagc	tgtttctgt	gtgaaattgt	tatccgctca	caattccaca	caacatacga	4320
gccggaagca	taaagtgtaa	agcctgggt	gcctaatgag	tgagctaact	cacattaatt	4380
gcgttgcgt	caactgcccgc	tttccagtcg	ggaaacctgt	cgtgccagct	gcattaatga	4440
atcgcccaac	gcccggggag	aggcggtttg	cgtattggc	gctttccgc	ttcctcgctc	4500
actgactcgc	tgcgctcgg	cggttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	4560
gtaatacggt	tatccacaga	atcagggat	aacgcaggaa	agaacatgt	agaaaaaggc	4620
cagaaaaagg	ccaggaaccg	taaaaaaggcc	gcgttgcgt	cgttttcca	taggctccgc	4680
ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtacga	ggtggcgaaa	cccgacagga	4740
ctataaagat	accaggcggt	tcccccttgg	agctccctcg	tgcgctctcc	tgttccgacc	4800
ctgcccctta	ccggataacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctca	4860
tgtcacgct	gttaggtatct	cagttcggtt	taggtcgttc	gctcaagct	ggctgtgt	4920
cacgaacccc	ccgttcagcc	cgaccgctgc	gccttacccg	gtaactatcg	tctttagtcc	4980
aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	5040
gcgaggtatg	taggcgggtc	tacagagttc	ttgaagtgg	ggcctaacta	cgctacact	5100
agaaggacag	tatttgtat	ctgcgctctg	ctgaagccag	ttaccttccg	aaaaagagtt	5160
ggtagcttt	gatccggcaa	acaaaccacc	gctggtagcg	gtggttttt	tgttgcaag	5220
cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	5280
tctgacgctc	agtggaaacga	aaactcacgt	taagggattt	tggcatgag	attatcaaaa	5340
aggatcttca	cctagatcct	tttaaattaa	aaatgaagtt	ttaaatcaat	ctaaagtata	5400
tatgagtaaa	cttggctctga	cagttaccaa	tgcttaatca	gtgaggcacc	tatctcagcg	5460
atctgtctat	ttcggtcatac	catagttgcc	tgactcccg	tcgtgtatg	aactacgata	5520
cgggagggtct	taccatctgg	ccccagtgt	gcaatgatac	cgcgagaccc	acgctcaccc	5580
gctccagatt	tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgcag	aagtggtcct	5640
gcaactttat	ccgcctccat	ccagtcattt	aattgttgc	gggaagctag	agtaagttagt	5700

tcgccagtt	atagttgcg	caacgttgg	gccattgcta	caggcatcgt	ggtgtcacgc	5760
tcgtcgttt	gtatggctt	attcagctcc	ggttccaaac	gatcaaggcg	agttacatga	5820
tccccatgt	tgtcaaaaaa	agcggttagc	tccttcggtc	ctccgatcgt	tgtcagaagt	5880
aagttggccg	cagtgttata	actcatgggt	atggcagcac	tgcataattc	tcttactgtc	5940
atgccatccg	taagatgctt	ttctgtgact	ggtgagtaact	caaccaagtc	attctgagaa	6000
tagtgtatgc	ggcgaccgag	ttgtcttgc	ccggcgtcaa	tacgggataa	taccgcgc	6060
catagcagaa	ctttaaaagt	gctcatcatt	ggaaaacgtt	tttcggggcg	aaaactctca	6120
agatatccat	cgctgttgcg	atccagttcg	atgtaaccca	ctcgtgcacc	caactgatct	6180
tcagcatctt	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	gaaaaatgcc	6240
gcaaaaaagg	gaataaggc	gacacggaaa	tgttgaatac	tcataacttt	ccctttcaa	6300
tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	6360
tagaaaaata	aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	acctgacgtc	6420

2020 TOE = 00000000000000000000000000000000

```

<210> 3
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker

<400> 3
Leu Ala Arg Leu Leu Ala Arg Leu Leu Ala Arg Leu
    1           5           10

```

```

<210> 4
<211> 67
<212> PRT
<213> Artificial sequence

<220>
<223> Conserved sequence of steroid/thyroid hormone
      receptor superfamily DNA-binding domain

```

```

<221> VARIANT
<222> (0)...(0)
<223> Xaa is non-conserved amino acids within the
      DNA-binding domain

<221> VARIANT
<222> 7, 9, 11, 13, 22, 27, 58, 61, 66
<223> amino acid residues that are almost universally
      conserved, but for which variations have been
      found in some identified hormone receptors

```

```

<400> 4
Cys Xaa Xaa Cys Xaa Xaa Asp Xaa Ala Xaa Gly Xaa Tyr Xaa Xaa Xaa
    1           5           10           15
Xaa Cys Xaa Xaa Cys Lys Xaa Phe Phe Xaa Arg Xaa Xaa Xaa Xaa Xaa Xaa
    20          25          30
Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Lys
    35          40          45
Xaa Xaa Arg Xaa Xaa Cys Xaa Xaa Cys Arg Xaa Xaa Lys Cys Xaa Xaa
    50          55          60
Xaa Gly Met
    65

```

```

<210> 5
<211> 34
<212> DNA

```

<213> Bacteriophage P1

<400> 5

ataaacttcgt atagcataca ttatacgaag ttat

34

<210> 6

<211> 34

<212> DNA

<213> E. coli

<400> 6

ataaacttcgt atagtataca ttatacgaag ttat

34

<210> 7

<211> 34

<212> DNA

<213> E. coli

<400> 7

acaacttcgt ataatgtatg ctatacgaag ttat

34

<210> 8

<211> 34

<212> DNA

<213> Saccharomyces cerevisiae

<400> 8

gaagttcccta ttctctagaa agtataaggaa cttc

34

<210> 9

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> peptide linker moiety

<400> 9

Gly Gly Gly Gly Ser

1

5

<210> 10

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety; sequence can be repeated  
indefinit number of times

<400> 10

Gly Gly Gly Gly Ser

1

5

<210> 11

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

2020/07/26 10:00:00

2020 RELEASE UNDER E.O. 14176

<400> 11  
Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Ser  
1 5 10

<210> 12  
<211> 14  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<400> 12  
Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly  
1 5 10

<210> 13  
<211> 18  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<400> 13  
Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Ser Gly Ser Thr  
1 5 10 15  
Lys Gly

<210> 14  
<211> 14  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<400> 14  
Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly  
1 5 10

<210> 15  
<211> 18  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<400> 15  
Gly Ser Thr Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser Thr  
1 5 10 15  
Lys Gly

<210> 16  
<211> 14  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<400> 16  
Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Glu Phe  
1 5 10

<210> 17  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<400> 17  
Ser Arg Ser Ser Gly  
1 5

<210> 18  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<400> 18  
Ser Gly Ser Ser Cys  
1 5

<210> 19  
<211> 27  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<400> 19  
Met Gly Arg Ser Gly Gly Cys Ala Gly Asn Arg Val Gly Ser Ser  
1 5 10 15  
Leu Ser Cys Gly Gly Leu Asn Leu Gln Ala Met  
20 25

<210> 20  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide linker moiety

<221> VARIANT  
<222> (0)...(0)  
<223> Xaa is (GmS)<sub>n</sub>, where m is  
from 2 to 4 and n is from 1 to 11.

<400> 20  
Ala Met Xaa Ala Met

1

5

<210> 21  
<211> 16  
<212> PRT  
<213> Drosophila acanthoptera

<400> 21  
Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys  
1 5 10 15